Host-associated differentiation and evidence for sexual reproduction in Iranian populations of the cotton aphid, Aphis gossypii

Razmjou, J ; Vorburger, C ; Moharramipour, S ; Mirhoseini, S Z ; Fathipour, Y

Abstract: Phytophagous insects with wide host ranges often exhibit host-associated genetic structure. We used microsatellite analysis to assess the population structure of the cotton aphid, Aphis gossypii Glover (Hemiptera: Aphididae), a serious pest on many economically important crops worldwide. We sampled aphids from five host plant species in Iran and detected strong population subdivision, with an overall multilocus FST of 0.191. The matrix of pairwise FST values indicated that differentiation between populations collected from different hosts was significantly stronger than between populations from the same hosts. Host-associated differentiation was further supported by Bayesian clustering analyses, which grouped all samples from cotton together with aubergine, and all samples from cucumber together with pumpkin and hibiscus. This adds to the growing body of evidence that many seemingly generalist aphids are in fact an assemblage of host-specialized lineages. Although we detected a clear genetic signature of clonal reproduction, the genotypic diversity of A. gossypii in Iran is much higher than in other parts of the world. Particularly samples from cotton exhibited a surprisingly high genotypic diversity, suggesting that many lineages on this host are cyclical parthenogens that engage in regular bouts of sexual reproduction.

DOI: https://doi.org/10.1111/j.1570-7458.2009.00951.x

Host-associated differentiation and evidence for sexual reproduction in Iranian populations of the cotton aphid, *Aphis gossypii* Glover (Homoptera: Aphididae)

Jabraeil Razmjou[1,2], Christoph Vorburger[3] , Saeid Moharramipour[1,*], Seyed Ziaeddin Mirhoseini[4] and Yaghoub Fathipour[1]

[1]Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran

[2]Current address: Faculty of Agriculture, University of Mohaghegh Ardabili, Ardabil, Iran

[3]Institute of Zoology, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

[4]Faculty of Natural Resources, Guilan University, Someh Sara, Iran

Running title: Differentiation in Iranian populations of *Aphis gossypii*

*correspondence:

Saeid Moharramipour, Department of Entomology, Faculty of Agriculture, Tarbiat Modares University, PO Box 14115-336, Tehran, Iran.

E-mail: moharrami@modares.ac.ir

Abstract

Phytophagous insects with wide host ranges often exhibit host-associated genetic structure. The cotton aphid (*Aphis gossypii*) - a serious pest on many economically important crops worldwide - was proposed to be such a case. We used microsatellite analysis to assess the population structure of *A. gossypii* in Iran, including samples from five different host plants. We detected strong population subdivision with an overall multilocus F_{ST} of 0.191. The matrix of pairwise F_{ST} values indicated that differentiation between populations collected from different hosts was significantly stronger than between populations from the same hosts. Host-associated differentiation was further supported by Bayesian clustering analyses, which grouped all samples from cotton together with aubergine, and all samples from cucumber together with pumpkin and hibiscus. This adds to the growing body of evidence that many seemingly generalist aphids are in fact an assemblage of host-specialized lineages. Although we detected a clear genetic signature of clonal reproduction, the genotypic diversity of *A. gossypii* in Iran is much higher than in other parts of the world. Particularly samples from cotton exhibited a surprisingly high genotypic diversity, suggesting that many lineages on this host are cyclical parthenogens that engage in regular bouts of sexual reproduction.

Keywords: *Aphis gossypii*, cyclical parthenogenesis, host specialization, microsatellites, phytophagous insects

Introduction

Phytophagous insects are often characterized by a highly specialized resource use. Yet the diet breadth within clades of phytophagous insects varies considerably and generalist species are by no means uncommon. In several cases, however, reciprocal transplant experiments or the application of population genetic markers revealed that presumed generalist species actually consisted of a number of host-adapted subpopulations with limited genetic exchange, typically referred to as host races (e.g. McPheron et al., 1988; Waring et al., 1990; Via, 1991; Emelianov et al., 1995). Such findings suggest that host-based diversification is ongoing in these species and that host race formation may represent a first step in ecological speciation (Berlocher & Feder, 2002; Dres & Mallet, 2002).

Intraspecific host specialization is also known from a number of aphid species, e.g. the pea aphid, *Acyrthosiphon pisum* (Via, 1991; Frantz et al., 2006; Ferrari et al., 2008) or the black bean aphid, *Aphis fabae* (Mackenzie, 1996; Raymond et al., 2001). The aphids' ancestral reproductive mode is cyclical parthenogenesis, with many asexual generations of viviparous females during spring and summer, followed by a single sexual generation of males and egg-laying females in autumn, which produce the overwintering eggs. The switch between sexual and asexual reproduction is associated with host alternation in several aphid species. Mating and egg laying take place on the primary host, typically a woody plant, from which the first parthenogenetic generations disperse to herbacous secondary hosts. Host specialization on different secondary hosts is especially intriguing in host-alternating species like *A. fabae*, for example, because all lineages return to the same primary host in autum and presumably interbreed. However, many aphid species show numerous and irreversible transitions from cyclical to obligate

parthenogenesis, especially in warmer parts of the world (Simon et al., 2002), which entails that their life-cycle no longer requires a primary host.

Aphids are severe pests of agriculture, horticulture and forestry (Blackman & Eastop, 2000). The cotton aphid, *Aphis gossypii* Glover (Homoptera: Aphididae), is a polyphagous species with a worldwide distribution. It is a major pest of many important crops, including cotton, cucurbits, citrus, aubergine, potato, okra and many ornamental plants. It causes the greatest problems in European cucumber greenhouses (van Steenis & El-Khawass, 1995; Blackman & Eastop, 2000). In addition to damage caused by feeding and honeydew production, *A. gossypii* also transmits more than 50 plant viruses (Ebert & Cartwright, 1997; Blackman & Eastop, 2000). In Iran, *A. gossypii* is an important pest of cotton, cucurbits and greenhouse plants (Mojeni & Rezvani, 1996; Razmjou et al., 2006; Zamani et al., 2006). The taxonomic status of *A. gossypii* is problematic. It is highly variable not only in its morphology, but also in its life cycle and ecological characteristics (Blackman & Eastop, 2000). For example, temperature-dependent development and fecundity differ between lines collected from different host plants and geographic regions (van Steenis & El-Khwass, 1995; Kersting et al., 1999; Razmjou et al., 2006; Zamani et al., 2006). In addition, several studies provided strong evidence that genetically distinct host races exist in *A. gossypii* (Vanlerberghe-Masutti & Chavigny, 1998; Blackman & Eastop, 2000). Distinct forms of *A. gossypii* are reported from chrysanthemum and cucumber in Europe, with clones collected from cucumber performing poorly on chrysanthemum and vice versa (Guldemond et al., 1994). Of particular interest are two recent genetic surveys of *A. gossypii* populations in North Cameroon and Tunisia, employing microsatellites (Brévault et al., 2008; Charaabi et al., 2008). Both studies revealed an extremely low genotypic diversity, consistent with obligate parthenogenesis in this species, and strong host specialization with

specific clonal genotypes being associated with certain host plants or plant families. To assess whether obligate asexuality and strong host-specialization at the level of individual genotypes also characterise *A. gossypii* in other parts of its distribution, we studied the genetic diversity of this pest from different host plants and several regions of Iran.

Materials and Methods

Sample collection

Samples of *A. gossypii* used in this study were collected in Iran during August and September 2003 and 2004. Eighteen to 28 individuals were obtained from each location and host plant. Overall, we sampled 11 populations of *A. gossypii* on the basis of location/host plant (Fig. 1; Table 1). Aphids from the same individual plant were considered one sample and only one individual per sample was analyzed. Samples were taken no closer than one meter apart to avoid sampling the offspring of a single female. The samples were stored at -20°C or preserved in 75% ethanol and kept at 4°C prior to DNA extraction. In total, 245 individuals were genotyped.

DNA Extraction

DNA isolation was performed following the Chelex (5%) method according to Vanlerberghe-Masutti & Chavigny (1998). Aphids were crushed with a sterilized pipette in 1.5 mL Eppendorf tubes before 200 µL of a 5% (w/v) Chelex resin solution (Bio-Rad Laboratories) were added. The tube was heated to 56°C for 30 min, then 96°C for 5 min, vortexed and centrifuged for a few seconds. The supernatant was diluted 10 times and stored at -20°C until use as template DNA in PCRs.

Microsatellite analysis

Our samples were genotyped at four microsatellite loci: Ago53, Ago59, Ago66 and Ago69 (Vanlerberghe-Masutti et al., 1999). PCR reactions were performed in 15 µL volumes containing 0.6 units of Taq polymerase, 1x Mg^{2+} free reaction buffer, 1.5 mM $MgCl_2$, 200 µM dNTPs (Promega), 18 pmol of each primer and 1.5 µL of aphid template DNA (approximately 10 ng). Cycling conditions were as follows: initial denaturation at 94°C for 5 min; followed by 35 cycles of denaturation at 94°C for 1 min, locus-specific annealing temperature (Ago69: 65°C; other loci: 67°C) for 1 min and elongation at 74°C for 30 s; and a final extension step at 74°C for 5 min. PCR products were denatured at 94°C for 3 min and separated on 6% polyacrylamide urea gel at 75 Watt constant power using a sequencing apparatus (Bio-Rad Laboratories). After electrophoresis, PCR products were visualized by silver-staining using a procedure as described by Promega (1993). Allele sizes were determined with reference to ladders V and VIII by Roche Ltd.

Genetic analyses

We used the software FSTAT 2.9.3 (Goudet, 2001) to calculate expected and observed heterozygosities and test for deviations from linkage and Hardy-Weinberg equilibria. Because the clonal amplification of genotypes inevitably leads to deviations from genetic equilibria (Sunnucks et al., 1997, Halkett et al., 2005), these analyses were done without clonal copies, i.e. with the data reduced to a single representative of each multilocus genotype per population. *F*-statistics (Weir & Cockerham, 1984) were also calculated with FSTAT. To test for the possible occurrence of null alleles at the microsatellite loci, we applied the software MICRO-CHECKER

(van Oosterhout et al., 2004). The programs GENOTYPE and GENODIVE (Meirmans & van Tienderen, 2004) were used to identify identical genotypes and calculate clonal diversity statistics.

To explore our data for genetic signatures of host specialization, we used three different approaches. First, we estimated the relative contributions of the factors sampling site and host plant to the observed variance in allele frequencies with a hierarchical analysis of molecular variance (AMOVA) using ARLEQUIN 3.1 (Excoffier et al., 2005). Then we tested whether genetic differentiation was stronger between samples collected from different host plants than between samples from the same plant. For this we used a partial Mantel test implemented in the software ZT (Bonnet & Van de Peer, 2002), comparing the matrix of pairwise differentiation expressed as $F_{ST}/(1 - F_{ST})$ (Rousset, 1997) with a matrix expressing whether two samples were from the same or different plants, while controlling for the effect of geographic distances.

Finally, we inferred population structure without prior knowledge of the genotypes' site- and host-association using a Bayesian clustering algorithm as implemented in the software STRUCTURE 2.1 (Pritchard et al., 2000, Falush et al., 2003). This method assumes that within populations, loci are at Hardy-Weinberg and linkage equilibrium. However, it was shown that it allows meaningful inference even when applied to organisms with clonal or partially clonal reproduction, where these assumptions are frequently violated (e.g. Halkett et al., 2005; Delmotte et al., 2008). Carlsson (2008) has further shown that microsatellite null alleles have very small effects on the accuracy of assignment and on conclusions regarding the presence or absence of genetic differentiation based on this method. For all simulations, the admixture model and uninformative priors were used. We varied the number of genetic clusters (K) from 1 to 11 and ran five independent simulations for each K with a burn-in period of 50'000 iterations

followed by another 50'000 iterations. The most probable number of genetic clusters based on the log probability of the data was inferred following the method of Evanno et al. (2005). Based on a histogram of assignment probabilities, we decided to consider a genotype as assigned to a single cluster if its assignment probability to that cluster was greater than 80%. To assess the robustness of the results from our relatively short simulations we also ran five independent simulations with 750'000 MCMC steps after a burn-in of 500'000 iterations for the more likely values of K = 1-5. The results were extremely similar. We used the assignment probabilities form the longer runs for the graphical illustration of the results.

Results

Genic and genotypic diversity

The number of alleles detected at the four microsatellite loci ranged between two and six (Table 2). Among the 245 individuals, a total of 118 different multilocus genotypes could be distinguished. We refer to these as clones, although we acknowledge that with the limited resolution provided by the four loci, we are likely to underestimate the true number of different clones. Without clonal copies, i.e. with population samples reduced to a single representative of each clone, we detected no significant deviations from linkage equilibrium between any of the loci. However, some marked deviations from Hardy-Weinberg equilibrium were observed. Ago53 and Ago66 exhibited excess heterozygosity in most populations, which is reflected in negative values for overall F_{IS} (Table 2), whereas Ago69 exhibited significant homozygote excess (Table 2). Heterozygote excess is commonly found in aphid populations that contain a substantial fraction of obligately parthenogenetic genotypes (e.g. Delmotte et al., 2002; Vorburger et al., 2003; Halkett et al., 2005). It appears to be a consequence of the accumulation

of microsatellite mutations in lines that undergo long-term parthenogenesis (Wilson et al., 2003). Excess homozygosity is unusual, however, and may be caused by null alleles. Indeed, the analysis with MICRO-CHECKER provided evidence for the presence of null alleles at locus Ago69 in two populations, FCu and VC (Table 1). We therefore adjusted the allele frequencies at this locus according to van Oosterhout's correction provided by the software and re-ran several analyses to assess whether null alleles at this locus affected any of our conclusions.

Host-associated population structure

The overall multilocus F_{ST} value of 0.191 (0.193 after adjusting for null alleles) is highly significant and indicates substantial population structure. The matrix of pairwise F_{ST} values shows high values particularly among populations collected from different hosts (Table 3). Indeed, the partial Mantel test indicates that population samples from different hosts are more differentiated than samples from the same host plant (r = 0.359, P = 0.006), a result that is virtually unchanged by correction for null alleles (r = 0.346, P = 0.007). Host-based genetic differentiation is further supported by AMOVA, showing that host plant nested within site explains 18.8% of the the variance in allele frequencies, whereas site only explains 0.4% (site: VC = 0.005, df = 6, P = 0.379; plant within site: VC = 0.220, df = 4, P < 0.001; within plant: VC = 0.944, df = 271, P < 0.001). However, this analysis should be interpreted with caution because it partially confounds geographic and host-based variation, as several samples represented just a single host from a single site (see Table 1).

The distribution of the log-likelihoods for the number of genetic clusters (K) from the Bayesian clustering analysis with STRUCTURE peaked at estimates of K = 2 – 3. However, with K = 3, only 66.0% of individuals could be assigned to one of the clusters with more than 80%

probability, whereas 86.5% could be assigned with $K = 2$. The ΔK method by Evanno et al. (2005) also favoured $K = 2$ over $K = 3$ with a seven-fold higher value of ΔK. The assignment to these two clusters with respect to host plant is illustrated in Figure 2a: Most genotypes collected from cucumber, pumpkin and hibiscus were assigned to cluster 2, while all genotypes from aubergine and most genotypes from cotton were assigned to cluster 1. Figure 2 further shows that only individuals from cotton comprised a substantial fraction of genotypes that could not be assigned confidently to one of the two clusters, and that samples from cucumber, pumpkin and hibiscus each contained one or two apparent migrants, i.e. genotypes assigned to cluster 1. Under $K = 3$ (Fig. 2b), individuals from aubergine, cucumber, pumpkin and hibiscus remained well-defined groups, with those from aubergine assigned to cluster 1, and those from the latter three hosts assigned to cluster 3 (except the putative migrants). Genotypes assigned with high probability to cluster 2 were only found on cotton, yet overall, individuals from cotton represented a poorly defined group under $K = 3$.

Distribution of multilocus genotypes

Of the 118 different multilocus genotypes (MLGs) that could be distinguished, 35 were collected more than once. Most of those genotypes were only locally abundant, although 11 were collected at more than one sampling site. The suspected host specialization was also reflected by the common genotypes. Only nine genotypes were collected on more than one host, and in only two of those cases did this include hosts from both putative groups of plants indicated by the clustering analyses. Genotype nr. 1 was collected from aubergine as well as cucumber, and genotype nr. 27 was found on cotton, cucumber and pumpkin. The distribution of the most common genotypes (collected ≥ 5 times) among our samples is detailed in Table 4.

Interestingly, the clonal diversity appeared to differ among hosts. With the exception of site Kashmar, samples from cotton were very diverse, containing nearly as many genotypes as individuals (Table 1). Samples from other hosts comprised many copies of the same multilocus genotypes. The simplest measure of clonal diversity is the *G*/*N* ratio, the number of genotypes divided by the number of individuals (Table 1). Comparing this measure between cotton and cucumber, for which we had four samples each, there was indeed a significant difference (Mann-Whitney U-test, $P = 0.029$).

Discussion

Our microsatellite analysis detected strong population structure in Iranian *A. gossypii*, with a clear genetic signature of host specialization. The main split appears to be between genotypes associated with aubergine and cotton and genotypes associated with cucurbits (cucumber and pumpkin) and hibiscus. We cannot exclude that there may be further substructure within the first group as the Bayesinan clustering analysis provided some evidence for a third genetic cluster associated with cotton.

Host-related genetic structure in general, and a genetic differentiation between aphids from cotton and cucurbits in particular, are consistent with results from Vanlerberghe-Masutti & Chavigny (1998). In their study using RAPD fingerprints, they detected genetic differentiation among cucurbit and non-cucurbit hosts and distinctively different profiles between *A. gossypii* from cotton and cucumber collected in Laos. This is further supported by a microsatellite study on *A. gossypii* in northern Cameroon, where cotton and cucurbits were found to be colonized by distinct groups of clonal genotypes (Brévault et al., 2008). A similar study on *A. gossypii* from

Tunisia suggested even finer diversification, as specific clonal genotypes were associated with each of four plant families studied (Charaabi et al., 2008). What is remarkably different in our study is the much higher genotypic diversity. Despite the limited resolution provided by only four microsatellite loci, we could distinguish 118 different multilocus genotypes among the 245 individuals analysed. With a higher resolution from using eight microsatellites, the studies by Brévault et al (2008) and Charaabi et al. (2008) each detected only 11 multilocus genotypes (one found in both studies) in large samples of 1176 and 559 individuals, respectively. A number of genetic surveys of several species of aphids have shown that such a low genotypic diversity is a characteristic of populations consisting of obligate parthenogens (Sunnucks et al.,. 1996; Fenton et al., 1998, Fuller et al., 1999; Wilson et al., 1999; Llewellyn et al., 2003, Vorburger et al., 2003). Samples from aphid populations that reproduce predominantly by cyclical parthenogenesis are more diverse and largely consist of unique genotypes, despite the fact that also cyclical parthenogens are clonal during the growth season (Delmotte et al., 2002; Wilson et al., 2002). *Aphis gossypii* was thought to reproduce by obligate parthenogenesis in most parts of the world (Blackman & Eastop, 2000), yet there are reports of sexual reproduction of this species from North America and East Asia, where it may utilise several unrelated plants as primary hosts, such as members of the genera *Catalpa*, *Celastrus*, *Hibiscus*, *Punica* and *Rhamnus* (summarised in Blackman & Eastop, 2007). Particularly interesting is a report of cyclical parthenogenesis without host alternation of this species on cotton in China (Zhang & Zhong, 1982). The comparatively high genotypic diversity detected in our study suggests that the Caspian Sea region of the Middle East should be added to the list of areas where cyclical parthenogenesis of *A. gossypii* may occur. Much of this diversity was found in samples from cotton, which were more diverse on average than samples from other hosts. The cotton sample

from Varamin, for example, contained as many genotypes as individuals (Table 1). This strongly indicates that cotton-infesting aphids in Iran do engage in sexual reproduction, i.e. comprise cyclically parthenogenetic lineages. Whether these lineages exhibit host-alternation, remains to be investigated.

Another unexpected result from our study is that genotypes from hibiscus mostly clustered with genotypes collected from cucurbits and differed from those on aubergine and cotton. This is interesting because hibiscus and cotton belong to the same plant family (Malvaceae). Hibiscus is one of the plants found to be used as a primary host by cyclical parthenogens in North America and East Asia (Blackman & Eastop, 2007). However, it is unlikely that the observed association of cucurbit- and hibiscus-feeding aphids in Iran can be explained by cucurbit-specialized aphids using hibiscus as a primary host. First of all, the samples from hibiscus were collected in mid-summer, when cyclical parthenogens of *A. gossypii* should already have dispersed to their secondary hosts. Secondly, population samples from cucurbits exhibited a rather low clonal diversity with many genotypes occurring in multiple copies (Table 1). Thus it seems that the aphids from hibiscus in the present study used hibiscus as a secondary host and belong to the same genetic cluster as those collected from cucurbits.

Aphids collected from cotton represented the genetically most heterogeneous group. They comprised at least a few individuals from all clusters and most of those genotypes that could not be confidently assigned in the Bayesian clustering analyses. It was reported that *A. gossypii* from cucumber could not be reared on chrysanthemum (a host we did not sample from), and vice versa, but that aphids from both of these hosts could be reared on cotton (Guldemond et al., 1994). Together, these findings suggest that cotton may be a host plant that is suitable for more than just one host race of *A. gossypii*, which could also explain the high proportion of apparently

admixed genotypes on cotton. A similar scenario has been suggested for the pea aphid, in which there are many host-specialized sub-populations that all perform well on the 'universal host' *Vicia faba* (Ferrari et al., 2008). This situation differs somewhat from that reported by Brévault et al. (2008), who found that cotton was only colonized by very few clones in Cameroon, yet this may simply reflect the much lower genotypic diversity there.

To summarize, we show strong host-associated genetic differentiation within the polyphagous aphid *A. gossypii*. This adds to the growing body of evidence that many seemingly generalist aphids are in fact an assemblage of host-specialized lineages (Via, 1991; Guldemond et al., 1994; Mackenzie, 1996; Sunnucks et al., 1997; Gorur et al., 2005; Frantz et al., 2006; Lozier et al., 2007), and it suggests that *A. gossypii* may be in the process of incipient speciation. We further provide genetic evidence for sexual reproduction of *A. gossypii* in Iran. The Middle East may thus represent a source of genotypic variation for this globally distributed pest.

Acknowledgements

We gratefully acknowledge B. Dalirsefat and Sakaki for their helps in the experiment. This work was conducted in genetic laboratory of Faculty of Natural Resources, Guilan University, Someh Sara, Iran. It was supported, in part, by Tarbiat Modares University. C.V. is supported by the Swiss National Science Foundation (grant 311A0-109266).

References

Berlocher SH & Feder JL (2002) Sympatric speciation in phytophagous insects: moving beyond controversy? Annual Review of Entomology 47: 773–815.

Blackman RL & Eastop VF (2000) Aphids on the World's Crops: an Identification and Information Guide, 2nd edn. John Wiley and Sons: Chichester.

Blackman RL & Eastop VF (2007) Aphids on the World's Herbaceous Plants and Shrubs: An Identification and Information Guide. John Wiley and Sons: Chichester.

Bonnet P & Van de Peer Y (2002) zt: a software tool for simple and partial Mantel tests. Journal of Statistical Software 7: 1-12.

Brévault T, Carletto J, Linderme D & Vanlerberghe-Masutti F (2008) Genetic diversity of the cotton aphid *Aphis gossypii* in the unstable environment of a cotton growing area. Agricultural and Forest Entomology 10: 215-223

Carlsson J (2008) Effects of microsatellite null alleles on assignment testing. Journal of Heredity 99: 616-623.

Charaabi K, Carletto J, Chavigny P, Marrakchi M, Makni M & Vanlerberghe-Masutti F (2008) Genotypic diversity of the cotton-melon aphid *Aphis gossypii* (Glover) in Tunisia is structured by host plants. Bulletin of Entomological Research 98: 333-341.

Delmotte F, Leterme N, Gauthier JP, Rispe C & Simon JC (2002) Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. Molecular Ecology 11: 711–723.

Delmotte F, Giresse X, Richard-Cervera S, M'Baya J, Vear F, Tourvieille J, Walser P & Tourvieille de Labrouhe D (2008) Single nucleotide polymorphisms reveal multiple

introductions into France of *Plasmopara halstedii*, the plant pathogen causing sunflower downy mildew. Infection, Genetics and Evolution 8: 534-540.

Dres M & Mallet J (2002) Host races in plant-feeding insects and their importance in sympatric speciation. Philosophical Transactions of the Royal Society of London B Biological Sciences: 471–492

Ebert TA & Cartwright B (1997) Biology and ecology of *Aphis gossypii* Glover (Homoptera: Aphididae). Southwestern Entomologist 22: 116-153.

Emelianov I, Mallet J & Baltensweiler W (1995) Genetic differentiation in *Zeiraphera diniana* (Lepidoptera: Tortricidae, the larch budmoth): polymorphism, host races or sibling species? Heredity 75: 416-424.

Evanno G, Regnaut S & Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611-2620

Excoffier L, Laval G & Schneider S (2005) Arlequin (version 3.0): An integrated software for population genetics data analysis. Evolutionary bioinformatics online 1: 47-50

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567–1587

Fenton B, Woodford JA, & Malloch G (1998) Analysis of clonal diversity of the peach-potato aphid, *Myzus persicae* (Sulzer), in Scotland, UK and evidence for the existence of a predominant clone. Molecular Ecology 7: 1475–1487.

Ferrari J, Via S, & Godfray HCJ (2008) Population differentiation and genetic variation in performance on eight hosts in the pea aphid complex. Evolution 62: 2508-2524.

Frantz A, Plantegenest M, Mieuzet L & Simon JC (2006) Ecological specialization correlates with genotypic differentiation in sympatric host-populations of the pea aphid. Journal of Evolutionary Biology19: 392-401.

Fuller SJ, Chavigny P, Lapchin L & Vanlerberghehe-Massutti F (1999) Variation in clonal diversity in glasshouse infestations of the aphid, *Aphis gossypii* Glover in southern France. Molecular Ecology 8: 1867–1877

Gorur G, Lomonaco C & Mackenzie A (2005) Phenotypic plasticity in host-plant specialization in *Aphis fabae*. Ecological Entomology 30: 657-664

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). http://www.unil.ch/izea/softwares/fstat.html.

Guldemond JA, Tigges WT & De Vrijer PWF (1994) Host races of *Aphis gossypii* (Homoptera: Aphididae) on cucumber and chrysanthemum. Population Ecology 23: 1235-1240.

Halkett F, Plantegenest M, Prunier-Leterme N, Mieuzet L, Delmotte F & Simon JC (2005) Admixed sexual and facultatively asexual aphid lineages at mating sites. Molecular Ecology 14: 325-336.

Kersting U, Satar S & Uygun N (1999) Effect of temperature on development rate and fecundity of apterus *Aphis gossypii* Glover reared on *Gossypium hirsutum* L. Journal of Applied Entomology 123: 23-27.

Llewellyn KS, Loxdale HS, Harrington R, Brookes CP, Clarke SJ & Sunnucks P (2003) Migration and genetic structure of the grain aphid (*Sitobion avenae*) in Britain related to climate and clonal fluctuation as revealed using microsatellites. Molecular Ecology 12: 21–34.

Lozier JD, Roderick GK & Mills NJ (2007) Genetic evidence from mitochondrial, nuclear, and endosymbiont markers for the evolution of host plant associated species in the aphid genus *Hyalopterus* (Hemiptera: Aphididae). Evolution 61(6): 1353–1367.

Mackenzie A (1996) A trade-off for host plant utilization in the black bean aphid, *Aphis fabae*. Evolution 50: 155-162.

McPheron BA, Smith DC & Berlocher SH (1988) Genetic differences between host races of *Rhagoletis pomonella*. Nature 336: 64-66.

Meirmans PG, van Tienderen PH (2004). GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Molecular Ecology Notes **4**: 792-794.

Mojeni DT, Rezvani A (1997) Biology and population fluctuating *A. gossypii* in the cotton fields of Gorgan. Journal of Entomological Society of Iran 16: 1-10.

Pritchard JK, Stephens M & Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Promega (1993) DNA silver staining system, Technical manual no. TMD005. Promega: Madison, WI, USA.

Raymond B, Seale JB, & Douglas AE (2001) On the processes shaping reproductive isolation in aphids of the *Aphis fabae* (Scop.) complex (Aphididae : Homoptera). Biological Journal of the Linnean Society 74: 205-215.

Razmjou J, Moharramipour S, Fathipour Y & Mirhoseini SZ (2006) Effect of cotton cultivar on performance of cotton aphid, *Aphis gossypii* Glover (Homoptera: Aphididae) in Iran. Journal of Economic Entomology 99: 1820-1824.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics 145: 1219-1228.

Simon JC, Rispe C & Sunnucks P (2002) Ecology and evolution of sex in aphids. Trends in Ecology & Evolution 17: 34-39.

Sunnucks P, De Barro PJ, Lushai G, Maclean N & Hales DF (1997) Genetic structure of an aphid studied using microsatellites: cyclic parthenogenesis, differentiated lineages, and host specialization. Molecular Ecology 6**:** 1059–1073.

Sunnucks P, England P, Taylor AC & Hales DF (1996) Microsatellite and chromosome evolution of parthenogenetic *Sitobion* aphids in Australia. Genetics 144: 747–756.

van Oosterhout C, Hutchinson WF, Wills DPM & Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535-538.

van Steenis MJ, El-Khawass KAMH (1995) Life history of *Aphis gossypii* on cucumber: influence of temperature, host plant and parasitism. Entomologia Experimentalis et Applicata 76: 121-131.

Vanlerberghe-Masutti F & Chavigny P (1998) Host-basic genetic differentiation in the aphid, *Aphis gossypii* Glover, evidenced from RAPD fingerprints. Molecular Ecology **7**: 905-914.

Vanlerberghe-Masutti F, Chavigny P & Fuller SJ (1999) Characterisation of microsatellite loci in the aphid, *Aphis gossypii* Glover. Molecular Ecology 8: 693–695.

Via S (1991) The genetic structure of host plant adaptation in a spatial patchwork - demographic variability among reciprocally transplanted pea aphid clones. Evolution 45: 827-852.

Vorburger C, Lancaster M & Sunnucks P (2003) environmentally related patterns of reproductive modes in the aphid *Myzus persicae* and the predominance of two 'superclones' in Victoria, Australia. Molecular Ecology 12: 3493-3504.

Waring GL, Abrahamson WG & Howard DJ (1990) Genetic differentiation among host-associated populations of the gall-maker *Eurosta solidaginis* (Diptera: Tephritidae). Evolution 44: 1648-1655.

Weir BS & Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. Evolution 38: 1358–1370.

Wilson ACC, Sunnucks P & Hales DF (1999) Microevolution, low clonal diversity and genetic affinities of parthenogenetic *Sitobion* aphids in New Zealand. Molecular Ecology 8**:** 1655 – 1666.

Wilson ACC, Sunnucks P, Blackman RL & Hales DF (2002) Microsatellite variation in cyclically parthenogenetic populations of *Myzus persicae* in south-eastern Australia. Heredity 88: 258–266.

Wilson ACC, Sunnucks P & Hales DF (2003) Heritable genetic variation and potential for adaptive evolution in asexual aphids (Aphidoidea). Biological Journal of the Linnean Society 79: 115-135.

Zamani, AA Talebi AA, Fathipour Y & Baniameri V (2006) Effect of temperature on biology and population growth parameters of *Aphis gossypii* Glover (Hom., Aphididae) on greenhouse cucumber. Journal of Applied Entomology 130 (8): 453–460.

Zhang GX & Zhong TS (1982) Experimental studies on some aphid life-cycle patterns. Sinozoologia 2: 2-17.

Table 1. Collection information for the 11 population samples included in this study. *N* = number of individuals genotyped, *G* = number of different multilocus genotypes in the sample.

Sample code	Location	Host plant	Collection date	*N*	*G*	*G/N* ratio
FAu	Faculty (Tehran)	aubergine	08.9.2004	20	6	0.30
FCu	Faculty (Tehran)	cucumber	08.9.2004	27	12	0.44
GC	Gorgan	cotton	19.8.2004	22	16	0.73
KC	Kashmar	cotton	18.8.2004	20	9	0.45
MC	Moghan	cotton	02.8.2005	25	24	0.96
MCu	Moghan	cucumber	02.8.2005	19	8	0.42
MH	Moghan	hibiscus	02.8.2005	19	12	0.63
MP	Moghan	pumpkin	02.8.2005	18	10	0.56
PCu	Pishva	cucumber	07.9.2004	22	6	0.27
SCu	Shahryar	cucumber	09.9.2004	25	9	0.36
VC	Varamin	cotton	07.9.2004	28	28	1.00

Table 2. Number of alleles, expected and observed heterozygosities and Weir and Cockerham's (1984) *F*-statistics for the four loci used in this study.

Locus	No. alleles	H_E	H_O	F_{IT}	F_{IS}	F_{ST}
Ago53	2	0.196	0.228	-0.037	-0.113	0.068***
Ago59	6	0.523	0.497	0.249***	0.094	0.171***
Ago66	4	0.486	0.628	-0.019	-0.271***	0.198***
Ago69	4	0.485	0.385	0.462***	0.286***	0.246***
overall	4	0.423	0.434	0.210***	0.024	0.191***

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$

Tab. 3. Pairwise F_{ST} values among 11 samples of *Aphis gossypii* from Iran, arranged by host plant.

		Cotton				Cucumber				Pumpkin	Aubergine
		Gorgan	Kashmar	Moghan	Varamin	Faculty	Moghan	Pishva	Shahryar	Moghan	Faculty
Cotton	Gorgan										
	Kashmar	0.177***									
	Moghan	0.057***	0.147***								
	Varamin	0.157***	0.178***	0.058***							
Cucumber	Faculty	0.288***	0.295***	0.211***	0.173***						
	Moghan	0.242***	0.445***	0.239***	0.275***	0.266***					
	Pishva	0.212***	0.356***	0.185***	0.132***	0.063^{NS}	0.121*				
	Shahryar	0.269***	0.302***	0.225***	0.138***	0.128*	0.228***	0.048^{NS}			
Pumpkin	Moghan	0.141***	0.362***	0.178***	0.209***	0.177**	0.022^{NS}	0.014^{NS}	0.181***		
Aubergine	Faculty	0.201***	0.199**	0.167***	0.244***	0.287***	0.525***	0.400**	0.461***	0.350***	
Hibiscus	Moghan	0.200***	0.305***	0.173***	0.187***	0.124***	0.122***	0.082**	0.198***	0.048^{NS}	0.279***

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$

Table 4. Distribution of multilocus genotypes that were collected at least five times among the 11 Iranian samples of *Aphis gossypii*.

Genotype numbers are as assigned by the software GENOTYPE.

		Cotton				Cucumber				Pumpkin	Aubergine	Hibiscus
Genotype nr.	Total *N*	Gorgan	Kashmar	Moghan	Varamin	Faculty	Moghan	Pishva	Shahryar	Moghan	Faculty	Moghan
7	16					1		13	2			
1	10					1					9	
8	10					6		3	1			
2	8		2								6	
10	8					8						
65	8						5			3		
89	8								8			
64	7						4			3		
34	6		6									
5	5	1	2								2	
9	5					2		2	1			
27	5	1					1			3		
69	5						1	2	2			
71	5											5

Figure captions

Fig 1. Map of Iran showing the collection sites of *A. gossypii* and the host plants sampled at the different sites.

Fig. 2. Bar plots illustrating the results from the Bayesian clustering analysis in STRUCTURE 2.1, using (a) $K = 2$ clusters or (b) $K = 3$ clusters. Each genotype is represented by a thin vertical bar with different colours representing the assignment probabilities to each of the clusters. The analysis was performed without clonal copies, i.e. with a data set reduced to a single representative of each multilocus genotype per population.

Figure 1

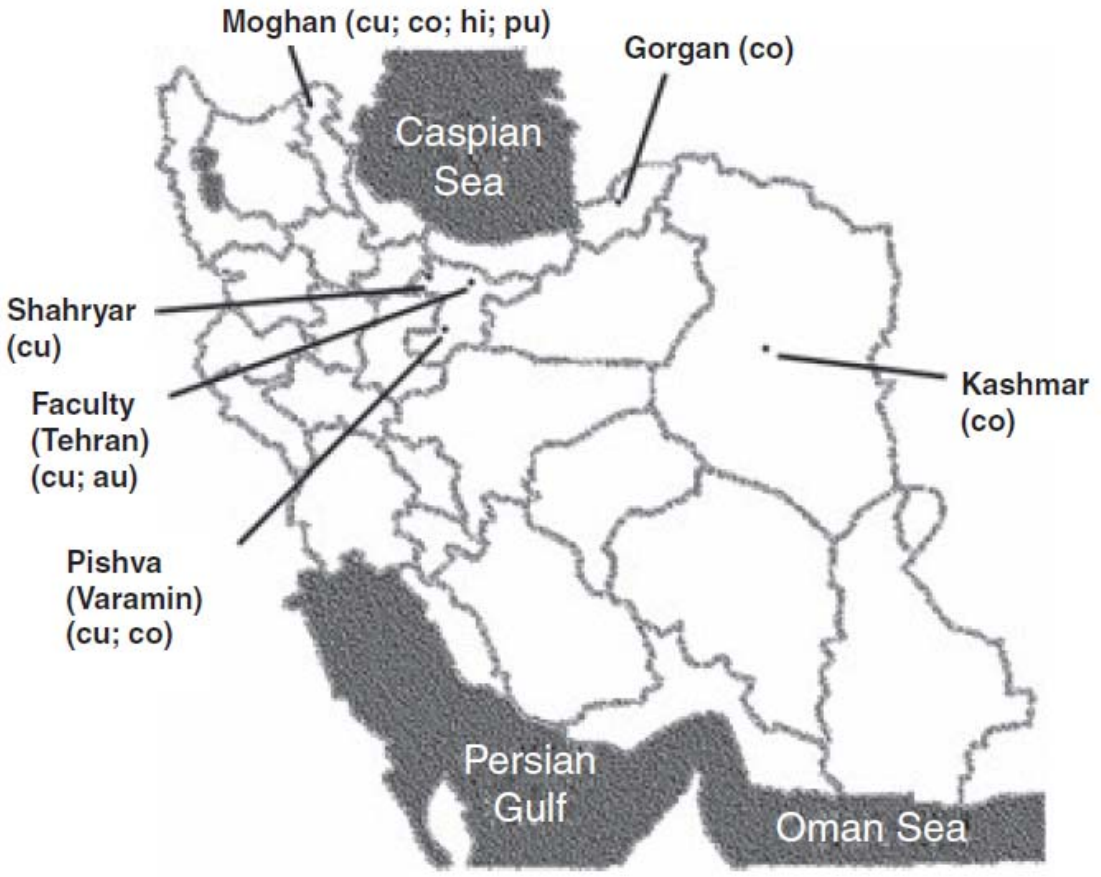
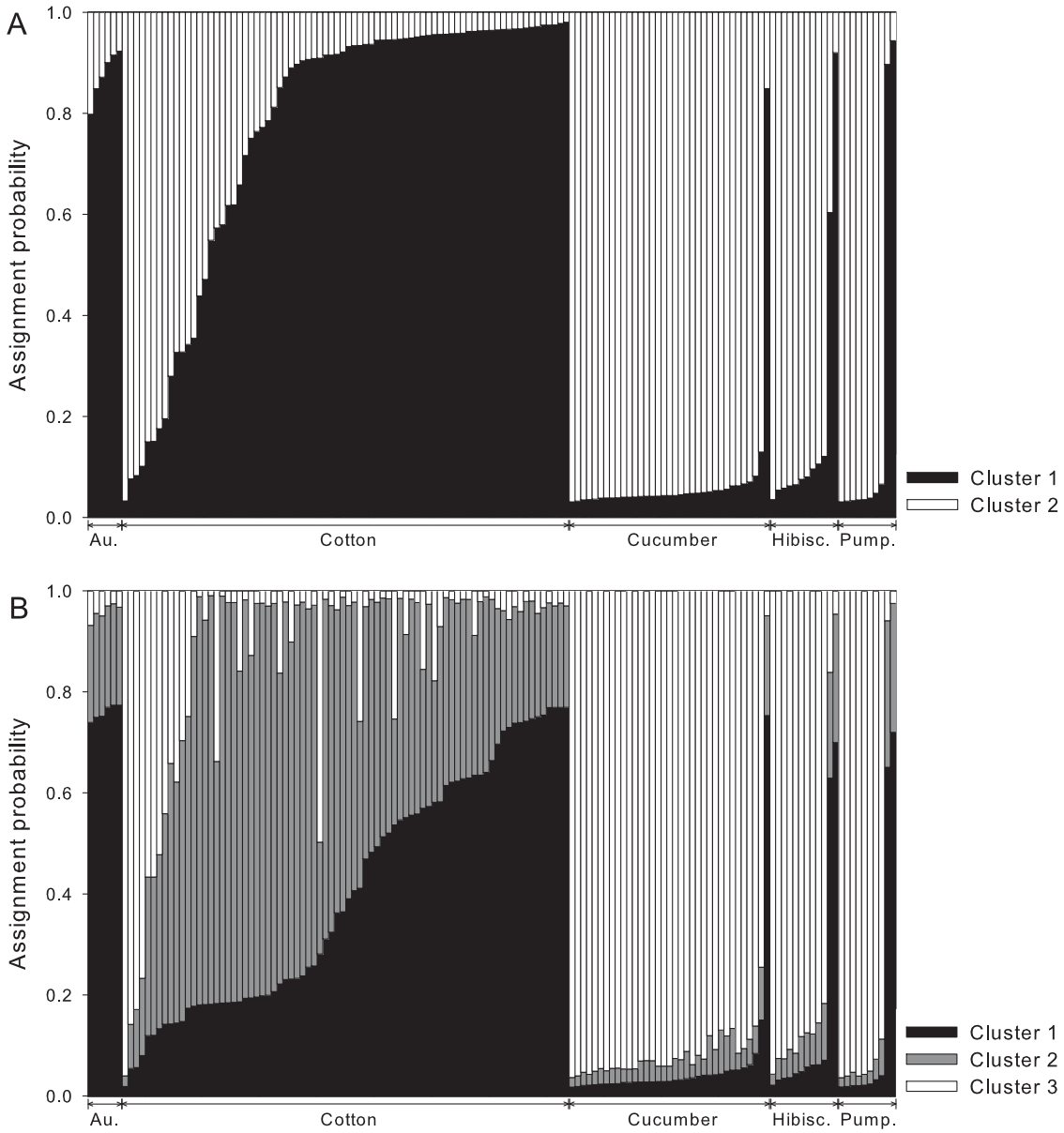

Moghan (cu; co; hi; pu)
Gorgan (co)
Caspian Sea
Shahryar (cu)
Faculty (Tehran) (cu; au)
Pishva (Varamin) (cu; co)
Kashmar (co)
Persian Gulf
Oman Sea

Figure 2


A
1.0
0.8
0.6
0.4
0.2
0.0
Assignment probability
Au.
Cotton
Cucumber
Hibisc.
Pump.
Cluster 1
Cluster 2
B
1.0
0.8
0.6
0.4
0.2
0.0
Assignment probability
Au.
Cotton
Cucumber
Hibisc.
Pump.
Cluster 1
Cluster 2
Cluster 3